

# High Sequence Conservation of Human Immunodeficiency Virus Type 1 Reverse Transcriptase under Drug Pressure despite the Continuous Appearance of Mutations

Francesca Ceccherini-Silberstein,<sup>1\*</sup> Federico Gago,<sup>2</sup> Maria Santoro,<sup>1</sup> Caterina Gori,<sup>3</sup>  
Valentina Svicher,<sup>1</sup> Fátima Rodríguez-Barrios,<sup>2</sup> Roberta d'Arrigo,<sup>3</sup>  
Massimo Ciccozzi,<sup>4</sup> Ada Bertoli,<sup>1</sup> Antonella d'Arminio Monforte,<sup>5</sup>  
Jan Balzarini,<sup>6</sup> Andrea Antinori,<sup>3</sup>  
and Carlo-Federico Perno<sup>3</sup>

*Department of Experimental Medicine, University of Rome "Tor Vergata," Rome,<sup>1</sup> National Institute for Infectious Diseases, "L. Spallanzani," Rome,<sup>3</sup> and Istituto Superiore di Sanità, Rome,<sup>4</sup> and Institute of Infectious and Tropical Diseases, University of Milan, Milan,<sup>5</sup> Italy; Department of Pharmacology, University of Alcalá, Alcalá de Henares, Spain<sup>2</sup>; and Rega Institute for Medical Research, K. U. Leuven, Leuven, Belgium<sup>6</sup>*

Received 22 March 2005/Accepted 19 May 2005

**To define the extent of sequence conservation in human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) in vivo, the first 320 amino acids of RT obtained from 2,236 plasma-derived samples from a well-defined cohort of 1,704 HIV-1-infected individuals (457 drug naïve and 1,247 drug treated) were analyzed and examined in structural terms. In naïve patients, 233 out of these 320 residues (73%) were conserved (<1% variability). The majority of invariant amino acids clustered into defined regions comprising between 5 and 29 consecutive residues. Of the nine longest invariant regions identified, some contained residues and domains critical for enzyme stability and function. In patients treated with RT inhibitors, despite profound drug pressure and the appearance of mutations primarily associated with resistance, 202 amino acids (63%) remained highly conserved and appeared mostly distributed in regions of variable length. This finding suggests that participation of consecutive residues in structural domains is strictly required for cooperative functions and sustainability of HIV-1 RT activity. Besides confirming the conservation of amino acids that are already known to be important for catalytic activity, stability of the heterodimer interface, and/or primer/template binding, the other 62 new invariable residues are now identified and mapped onto the three-dimensional structure of the enzyme. This new knowledge could be of help in the structure-based design of novel resistance-evading drugs.**

The activity of the human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT), essential for viral replication, is specifically required for the conversion of single-stranded genomic RNA into double-stranded viral DNA, which is later integrated into the host genomic DNA (25). For this reason, HIV-1 RT inhibitors (RTIs) are powerful inhibitors of HIV-1 replication and represent an important class of antiretroviral agents (16).

HIV-1 RT is a multifunctional enzyme that possesses RNA-dependent and DNA-dependent DNA polymerase activities as well as an RNase H activity that specifically degrades the RNA strand of RNA/DNA hybrids (25). Its structure is a heterodimer consisting of two subunits, termed p66 and p51. Although both subunits share the first 440 amino acids (aa) of the RT gene, their relative arrangements are significantly different, and the substrate-binding site is located only in the p66 subunit. p51 does not directly contribute to catalytic enzyme activities (25, 38) but is essential for loading the p66 subunit on

the template/primer (27). Indeed, the single p66 monomer is devoid of any catalytic activity.

As an intrinsic property, and in contrast to other DNA polymerases, HIV-1 RT lacks a proofreading function. This error-prone nature of RT, together with the high rate of virus production sustained by HIV-1 infection in vivo, highly contributes to the continuous generation of new viral variants (14, 48, 63). In addition, retroviral particles contain two copies of single-stranded RNA, and template switches occur frequently during reverse transcription, thus generating mutations and recombinations by intramolecular and intermolecular jumps. This intrinsic variability is dramatically enhanced under sub-optimal drug pressure, resulting in mutations particularly in the *pol* region encoding both RT and protease (PR) enzymes (13, 24, 35, 65).

The emergence of resistance is the inevitable consequence of incomplete suppression of HIV replication by the antiretroviral drugs and represents a frequent and major limitation of antiviral therapy (13, 55). As a consequence, sequencing of HIV-1 RT and PR has been recommended to guide the choice of antiretroviral therapy in clinical practice (17, 29). Most of the known RT inhibitor resistance mutations are in the 5' polymerase coding regions, particularly in the fingers and palm subdomains. Therefore, on a routine basis, only the first part of

\* Corresponding author. Mailing address: Department of Experimental Medicine, University of Rome Tor Vergata, Via Montpellier 1, 00133 Rome, Italy. Phone: 39-06-72596553. Fax: 39-06-72596039. E-mail: ceccherini@med.uniroma2.it.

the genetic information of RT (up to 240 to 300 aa) is generally analyzed (57).

To date, mutations at 61 residues in the RT gene have been related to treatment with experimentally tested RTIs (13, 24, 35, 55, 57). Of these, 18 sites are involved in resistance against the eight currently approved nucleoside and nucleotide RT inhibitors (NRTIs) and 16 sites are involved in resistance against the three currently approved nonnucleoside RT inhibitors (NNRTIs) (35) (Stanford HIV sequence database 2005 [http://hivdb.Stanford.edu]). This variability has been extensively studied (1, 6, 15, 19, 23, 24, 26, 30, 39, 46, 51, 52, 61, 67), yet limited information is available regarding the conserved regions of HIV RT (3, 34), even under strong antiviral pressure.

The identification of conserved regions within the HIV-1 genome allows a better characterization of essential regions in the viral proteins and can help in the design of new therapeutic strategies aimed to drive the virus to mutate at key amino acids that are crucial for the maintenance of sufficient viral fitness (54). The resulting virus may be replicating at such a limited rate that its ability to damage the immune system, and to further mutate to compensate for such decreased fitness, is minimal. The characterization of the conserved areas of HIV enzymes and proteins is therefore of crucial importance to investigate new and innovative attempts to control and prevent efficient virus spread in the body.

This research approach led us to investigate and define the conserved areas of HIV-1 PR (10). We then concentrated our attention on HIV-1 RT, in an attempt to define the minimal conserved structure of the first 320 aa of 2,236 HIV-1 RT sequences obtained from a well-defined cohort of 1,704 HIV-1-infected individuals (drug-naïve and drug-treated patients). Conserved amino acids have been analyzed and mapped onto the structure of both p66 and p51 subunits using available X-ray information. The identification of novel HIV-1 RT invariant regions can now be used in the rational design of new HIV-1 RT inhibitors with a much more favorable resistance profile.

## MATERIALS AND METHODS

**Patients.** The study included 2,236 sequences derived from 1,704 individuals infected with HIV-1 enrolled between 1998 and 2003 in the Italian Cohort of Antiretroviral Naïve patients or at different clinical centers in central Italy. Of these, 457 were naïve for antiretroviral drugs, and 1,247 were failing antiretroviral therapy regimens. Overall, drug-treated patients were experiencing a median of three highly active antiretroviral therapy regimens: 324 patients had been treated with two RTIs, 317 patients with three RTIs, 353 patients with four RTIs, and 778 patients with five or more RTIs; 100% of treated patients experienced at least one NRTI, 58.2% at least one NNRTI, and 86% at least one protease inhibitor (PI). Among NRTIs received, lamivudine was used in 91.3% of patients, zidovudine in 81.1%, stavudine in 77.3%, didanosine in 63.6%, abacavir in 21.2%, dideoxycytosine in 17.9%, and tenofovir in 2.1%; among NNRTIs, nevirapine was used in 36.3%, efavirenz in 31%, and delavirdin in 0.1% of the cases.

At the time of genotypic analysis, 99% of drug-treated patients were under treatment with NRTIs: 62.3% of patients were treated with lamivudine (median time, 582 days), 54.4% with stavudine (median time, 553 days), 35.8% with didanosine (median time, 432 days), 30.5% with zidovudine (median time, 577 days), 14.4% with abacavir (median time, 351 days), 2.6% with dideoxycytosine (median time, 457 days), and 1.7% with tenofovir (median time, 171 days). Among these RTI-treated patients, 35% were also under NNRTIs: 17.9% with nevirapine (median time, 470 days) and 17.1% with efavirenz (median time, 448 days). PR inhibitors were included in the last regimen in 1,103 patients (62%).

**HIV sequencing.** HIV genotyping was performed on plasma samples by means of a commercially available kit (the ViroSeq HIV-1 genotyping system; Celera

Diagnostics/Abbott Molecular Diagnostics, Rome, Italy) based on RNA extraction, retrotranscription, DNA amplification, and sequencing by an automated sequencer (ABI 3100; Applied Biosystems, Foster City, CA) (10, 47). The amplified product contained the entire PR (99-aa) open reading frame and the first 320 aa of the RT open reading frame.

Sequence data were analyzed by dedicated HIV genotyping system software that automatically assembles the seven sequence segments into a consensus sequence, which is then compared to an NL4.3 reference strain. For each individual, the quality endpoint was ensured by a coverage of the RT sequence by at least three sequence segments. Sequences having a mixture of wild-type and mutant residues at single positions were considered to have a mutation at that position. When the mixture contained two different mutations, both were considered and reported (see Fig. 2). HIV subtype analysis showed that 18 out of 457 (3.9%) drug-naïve patients carried non-clade B *pol* subtypes (F = 4, A = 2, C = 1, G = 1, and recombinant subtypes = 10). Similarly, 60 out of 1,247 RTI-treated patients (4.8%) carried non-clade B *pol* subtypes (F = 27, C = 11, G = 4, A = 4, D = 2, and recombinant subtypes = 12). Overall, the large majority of the sequences included in this analysis (2,102/2,236, 94%) carried a *pol* B subtype virus.

PR and RT genotypic mutations of HIV from each patient were recorded in an anonymous database that includes demographic, immunologic, virologic, and therapeutic parameters. The majority of nucleotide sequences from drug-naïve patients have been submitted previously to GenBank (47); see below for the accession numbers of the new sequences.

Published *pol* consensus sequences of pure HIV-1 subtypes (A, B, C, D, F1, F2, G, H, J, and K) (Los Alamos DB, <http://hiv-web.lanl.gov/content/hiv-db/mainpage.html>), together with GenBank (<http://www.ncbi.nlm.nih.gov>)-published *pol* lentiviral sequences (of HIV-2 P18042, visna/maedi virus AAM51650, feline immunodeficiency virus [FIV] AAM13444, equine infectious anemia virus [EIAV] AAK21112, caprine arthritis encephalitis virus [CAEV] AAG48629, ovine lentivirus AAO33139, puma lentivirus AAA67168, simian immunodeficiency virus [SIV] SIVS4 P12502 and SIVCZ P17283, and bovine leukemia virus [BLV] P19561), were examined in this study and aligned using the Clustal X sequence alignment program.

**Mutations.** Consensus B was used as a reference strain for the definition of mutations. Mutations associated (by in vitro and in vivo studies) with resistance to NRTIs and NNRTIs currently discovered (24, 35, 55), as well as mutations not yet associated with drug resistance, were analyzed. In particular, the frequency of mutations in the first 320 aa of RT was calculated and statistically compared using the  $\chi^2$  test (based on a 2-by-2 contingency table containing the numbers of isolates from untreated and treated persons and the number of isolates with and without mutations).

Amino acids not mutated or mutated with  $\leq 1\%$  prevalence over each cohort of patients (drug naïve or treated) were defined as conserved. HIV-1 RT depictions were colored according to mutation frequency rate, so that in specific figures blue corresponds to highly conserved residues (variability in  $\leq 1\%$  of patients), cyan to amino acids mutated only in 1 to 5% of patients, green to amino acids mutated in 5 to 10% of patients, yellow to amino acids mutated in 10 to 25% of patients, and red to residues that appear mutated in  $>25\%$  of patients (for further details, see figure legends).

Conservative (strongly or weakly) and nonconservative amino acid substitutions were recognized according to the ClustalW formalism (<http://www.ebi.ac.uk>). Residue substitutions were considered strongly conservative if amino acid alterations were from/to one of the amino acids within the following groups having similar biochemical characteristics: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, and FYW; they were weakly conservative if alterations were from/to one of the amino acids within the following groups having different biochemical characteristics: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEHQK, MILVF, and HFY.

**Structural analysis.** Several X-ray crystallographic structures of HIV-1 RT available from <http://www.rcsb.org/pdb> were used to map mutations and conserved residues in three dimensions, namely, HIV-1 RT in complex with efavirenz and solved at 2.5-Å resolution (PDB code 1FK9 [49]), the apoenzyme solved at 2.7-Å resolution (PDB code 1DLO [31]), and DNA-bound HIV-1 RT (PDB code 1RTD [32]). The public domain program ViewerLite 5.0 (Accelrys) was employed for visualization and display.

**Nucleotide sequence accession numbers.** Nucleotide sequences of isolates from patients treated with at least a PI within the highly active antiretroviral therapy regimens (10, 58) were recently submitted to GenBank with these accession numbers: AY855351 to AY855439, AY855441 to AY855458, AY855460 to AY855502, AY855504 to AY855556, AY855558 to AY855773, AY855775 to AY855795, AY855797 to AY855818, AY855820 to AY855836, and AY995408 to AY995555; the rest of the sequences are in the process of being submitted.

## RESULTS AND DISCUSSION

**Degree of HIV-1 RT conservation.** HIV-1 RT conservation *in vivo*, both in the absence and in the presence of antiviral pressure, was assessed by evaluating the first 320 aa of the HIV-1 RT in 457 *pol* sequences of drug-naïve patients and in 1,779 sequences of 1,247 RTI-treated patients. This comparative analysis allowed us to observe that HIV-1 RT *in vivo* requires the full preservation of at least two-thirds of its amino acids (some with still-unknown functions), and of large areas of its tertiary structure, to maintain a stable and functional enzyme.

(i) **Drug-naïve patients.** The analysis of sequences from drug-naïve patients showed conservation ( $\leq 1\%$  variability) in 233 out of 320 aa (73% overall conservation) (Fig. 1). A few invariant residues were scattered throughout the sequence either individually (K103, S105, N175, K201, E203, L282, G285, K287, and L295) or either pairs or triplets, whereas the majority (159/233, 68%) clustered into 14 regions comprising 5 to 29 consecutive invariant residues (Fig. 1). The nine longest invariant regions, containing some functionally important residues, were I (T7 to P19), II (K70 to K82), III (Q91 to P97), IV (T107 to S117, including the catalytic D110), V (F124 to S134), VI (N147 to P157), VII (Y181 to E194, including the catalytic D185 and D186), VIII (T216 to I244), and IX (W252 to Y271).

(ii) **Drug-treated patients.** In RTI-treated patients, despite the increased appearance of mutations associated with drug treatment, amino acid invariance was still highly maintained in 202 out of 320 aa (63% of conservation versus 73% in naïve patients,  $P = 0.278$ ). The majority (178/202, 88%) of these invariant residues clustered in pairs or triplets or in regions with more than four (up to 20) neighboring residues (Fig. 1), confirming that participation of contiguous residues in structural domains is optimally required for cooperative function and sustainability of enzyme activity. However, residues that clustered in regions comprising more than five consecutive invariant residues were strongly reduced (107/202, 53% versus 68% in drug-naïve isolates,  $P = 0.002$ ), and the number of single invariant residues in the group of RTI-treated patients was increased (from 2.8% to 7.5%;  $P = 0.018$ ).

Compared to the nine large conserved areas found in drug-naïve patients, the degree of conservation in drug-treated patients was almost identical at six regions (I, III, IV, V, VI, and IX). Minor exceptions were found at T7, K11, and F116 (with a variation frequency somewhat increased to  $\leq 3\%$  of patients) and V108 and Q151 (with a variation frequency increased up to 6% and 4% of patients, respectively). Interestingly, the decrease of variation of aa 146 (from 2 to 0.8%) that was found in the group of RTI-treated patients resulted in an upstream extension of conserved area VI up to position R143.

By contrast, a different picture, characterized by a substantial decrease of amino acid conservation, was observed in the other three large conserved areas (II, VII, and VIII). The conserved areas II and VII markedly shrank to smaller stretches, whereas the largest conserved region, VIII, was split into several isolated conserved residues and two smaller conserved areas (229 to 237 and 239 to 244) (Fig. 1).

These results are consistent with a previous analysis conducted on 1,210 different RT sequences from drug-naïve and NRTI-treated (but NNRTI-naïve) individuals (24). However,

this study addressed only the mutation frequency at RT positions 1 to 240; in addition, the focus was on characterizing RT mutations, while distinction between invariant residues (variability of  $< 1\%$ ) and residues with low variability ( $\leq 3\%$ ) was not made. Taking into account these considerations and despite a general agreement, differences with that analysis were found in the rate of conservation at a small number of amino acid positions (e.g., 36, 88, 172, and 175 and obviously NNRTI-associated residues [100, 181, 189, 190, 194, 225, and 227]) for which we found variability of  $> 1\%$  (and  $> 5\%$  at residues 181 and 190), not reported in the aforementioned study (24), and at residues 111, 115, 224, and 239, for which variability was reported to be  $> 3\%$  in NRTI-treated patients ( $< 1\%$  in our study). It is assumable that the differences in RT conservation/mutation patterns highlighted by these studies may reflect the circulation of distinct viruses in different geographic areas as well as variation in patients' characteristics and RTI regimens.

**Pattern of HIV-1 RT mutations.** We found several RT polymorphisms in drug-naïve patients. Among 87 variable residues, 42 were mutated in  $> 5\%$  of drug-naïve patients; of them, 14 were highly variable (replaced in  $> 25\%$ ) (Fig. 1). The frequency of the major drug resistance mutations in drug-naïve patients was consistently  $\leq 1\%$ , with only the exception of V118I (2.6%), while minor drug resistance mutations (or unknown mutations in drug resistance-associated residues) were present at low variability in drug-naïve patients: A98S (6.8%), K101R (2.2%), V106I (1.3%), E138A (3.5%), V179I (3%), and T215D (1.8%) (Fig. 2).

As expected, the development of drug resistance during therapy failure resulted in a general increase of mutations in the RT from RTI-treated patients, with 70 out of 320 residues (22%) found mutated in  $> 5\%$  of patients (compared to 42 in drug-naïve patients [13%],  $P = 0.005$ ). Of these, 20 were highly variable (replaced in  $> 25\%$ ) (Fig. 1).

Statistically significant differences ( $P < 0.05$  by  $\chi^2$  test) in the variability frequency between HIV-1 RT sequences from drug-naïve and RTI-treated patients were found at 53 out of 320 residues (17%) (shown in boldface type in Fig. 2). Among them, 17 known NRTI resistance sites (M41, E44, A62, K65, D67, T69, K70, L74, V75, F77, F116, V118, Q151, M184, L210, T215, and K219) and 16 known NNRTI resistance sites, not only for efavirenz and nevirapine (A98, L100, K101, K103, V106, V108, I135, E138, V179, Y181, Y188, V189, G190, P225, F227, and K238), were identified (35, 55, 57; <http://hivdb.stanford.edu>). Besides, seven new sites (E28, K122, G196, Q278, R284, T286, and A288) were positively associated with RTI treatment, together with nine sites (K20, T39, K43, E203, H208, D218, H221, K223, and L228) recently associated with NRTI therapy (24). Interestingly, at four new sites (T27, I50, R83, and V276) the variability frequency was statistically decreased in the group of RTI-treated patients relative to the group of drug-naïve patients ( $P < 0.004$ ; Fig. 2).

Regarding the type of mutations, some unusual changes at known NRTI resistance sites (less common than others at the same positions) (E44A, D67G, L74I, V75M, V75T, T215I, K219R, and K219N), as well as novel mutations not yet associated with resistance to any known RTI (K20R, E28A, E28K, V35M, T39A, K43E, K43Q, K43N, K122E, G196E, E203K, E203D, H208Y, D218E, H221Y, K223Q, K223E, L228H, L228R, Q278H, R284K, T286A, and A288T), were signifi-



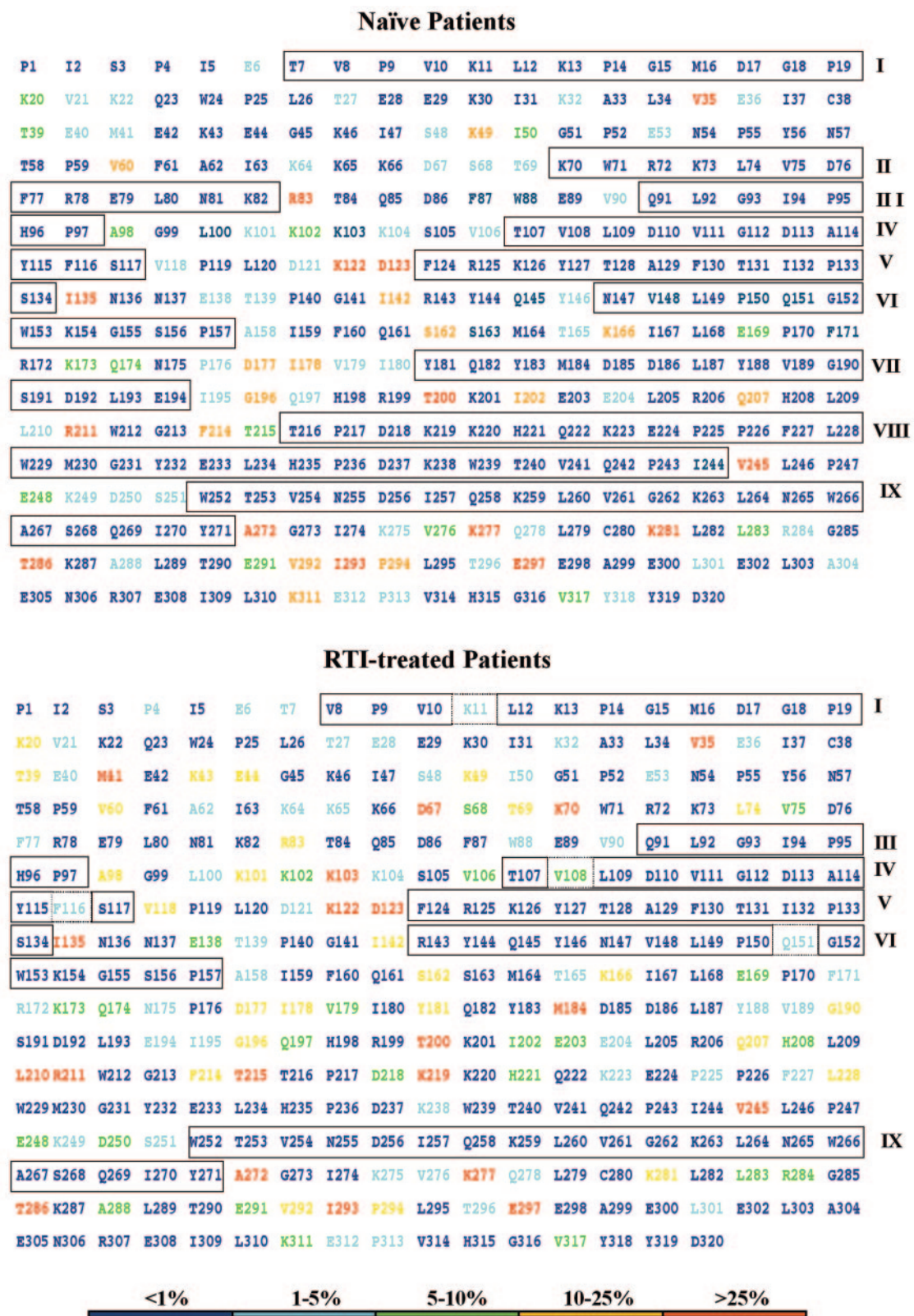


FIG. 1. Conserved regions of HIV-1 RT in drug-naïve and drug-treated HIV-1-infected patients. The sequence of the first 320 aa of HIV-1 RT of clade B consensus (shown as a reference) is colored according to the frequency rate of mutations observed in 457 plasma samples from drug-naïve patients and 1,779 plasma samples from RTI-treated patients. The long invariant regions are boxed (I to IX) in both drug-naïve and RTI-treated patients.

Naive	Treated	Naive	Treated	Naive	Treated
S <sub>1</sub>	P1 L <sub>2</sub> S <sub>1</sub>	Q <sub>1</sub> A <sub>2</sub>	P55 S <sub>3</sub> Q <sub>1</sub>	L109 I <sub>55</sub> M <sub>1</sub>	
I2	V <sub>2</sub> L <sub>1</sub>	F <sub>2</sub>	Y56 E <sub>1</sub> D <sub>1</sub>	G <sub>1</sub>	D110 N <sub>1</sub>
S3	C <sub>7</sub> N <sub>1</sub>		N57 I <sub>1</sub> H <sub>1</sub> Y <sub>1</sub>	L <sub>1</sub>	V111 I <sub>11</sub> A <sub>1</sub>
P4	S <sub>28</sub> T <sub>4</sub> L <sub>1</sub>	P <sub>1</sub> N <sub>1</sub> I <sub>1</sub> S <sub>1</sub>	T58 S <sub>2</sub> P <sub>1</sub> N <sub>1</sub>		G112 S <sub>5</sub> C <sub>1</sub>
T <sub>1</sub>	I5 V <sub>2</sub> T <sub>1</sub>		P59	Y <sub>1</sub>	D113 N <sub>2</sub> H <sub>2</sub> G <sub>1</sub>
G <sub>1</sub> K <sub>2</sub> D <sub>18</sub>	<u>E6</u> D <sub>33</sub> K <sub>13</sub> A <sub>3</sub> R <sub>1</sub> G <sub>1</sub>	M <sub>1</sub> I <sub>62</sub>	V60 I <sub>20</sub> M <sub>3</sub> L <sub>2</sub> E <sub>1</sub>		A114 V <sub>1</sub> S <sub>1</sub> E <sub>1</sub>
P <sub>3</sub>	T7 P <sub>6</sub> A <sub>7</sub> S <sub>2</sub> N <sub>1</sub> C <sub>1</sub>		F61		<u>Y115</u> F <sub>8</sub>
I <sub>3</sub>	V8 I <sub>5</sub> A <sub>2</sub>	V <sub>2</sub>	<u>A62</u> V <sub>66</sub> G <sub>2</sub> T <sub>1</sub> P <sub>1</sub>	L <sub>1</sub>	<b>F116</b> Y <sub>28</sub> S <sub>3</sub> H <sub>1</sub> L <sub>1</sub>
Q <sub>1</sub>	P9 Q <sub>1</sub> L <sub>1</sub> S <sub>1</sub>		I63		S117 A <sub>3</sub> L <sub>1</sub>
V10	A <sub>2</sub> L <sub>1</sub>	R <sub>9</sub>	K64 R <sub>39</sub> H <sub>12</sub> N <sub>6</sub> Y <sub>2</sub> E <sub>2</sub> Q <sub>1</sub>	G <sub>1</sub> D <sub>1</sub> I <sub>12</sub>	<b>Y118</b> I <sub>167</sub> M <sub>1</sub>
NR <sub>4</sub>	K11 R <sub>21</sub> T <sub>6</sub> Q <sub>2</sub> E <sub>2</sub> N <sub>1</sub> I <sub>1</sub> A <sub>1</sub>		<b>K65</b> R <sub>24</sub> E <sub>1</sub>		P119 L <sub>1</sub>
L12	F <sub>2</sub> I <sub>1</sub>	N <sub>1</sub>	K66 R <sub>3</sub> Q <sub>2</sub> N <sub>1</sub> E <sub>1</sub>		L120 S <sub>5</sub>
K13	R <sub>4</sub> E <sub>2</sub> Q <sub>1</sub> N <sub>1</sub> T <sub>1</sub> L <sub>1</sub>	A <sub>3</sub> N <sub>3</sub>	<b>D67</b> N <sub>670</sub> G <sub>28</sub> E <sub>14</sub> H <sub>4</sub> T <sub>2</sub> A <sub>3</sub> S <sub>1</sub> E <sub>1</sub>	G <sub>1</sub> Y <sub>6</sub> H <sub>10</sub>	D121 Y <sub>38</sub> C <sub>2</sub> N <sub>7</sub> H <sub>2</sub> E <sub>1</sub> A <sub>1</sub>
P14	A <sub>3</sub> S <sub>1</sub> L <sub>1</sub> T <sub>1</sub>	R <sub>1</sub> N <sub>1</sub> K <sub>1</sub> G <sub>17</sub>	S68 G <sub>69</sub> R <sub>3</sub> N <sub>2</sub> D <sub>1</sub>	H <sub>1</sub> T <sub>1</sub> N <sub>1</sub> V <sub>2</sub> Q <sub>2</sub> R <sub>4</sub> P <sub>14</sub> E <sub>100</sub>	<b>K122</b> E <sub>85</sub> P <sub>48</sub> R <sub>7</sub> A <sub>1</sub> Q <sub>4</sub> T <sub>2</sub>
G15	A <sub>1</sub>	P <sub>1</sub> D <sub>1</sub> A <sub>3</sub> N <sub>3</sub> S <sub>1</sub>	<b>T69*</b> D <sub>181</sub> N <sub>139</sub> S <sub>40</sub> A <sub>33</sub> I <sub>6</sub> G <sub>3</sub> P <sub>1</sub> R <sub>2</sub> E <sub>2</sub> INS <sub>9</sub>		D123 E <sub>91</sub> N <sub>48</sub> S <sub>28</sub> G <sub>12</sub> A <sub>3</sub> K <sub>4</sub> H <sub>1</sub>
V <sub>1</sub> T <sub>1</sub>	M16 V <sub>6</sub> F <sub>1</sub> I <sub>1</sub>	E <sub>2</sub> R <sub>4</sub>	<b>K70</b> R <sub>448</sub> E <sub>7</sub> S <sub>7</sub> G <sub>3</sub> N <sub>2</sub> D <sub>2</sub>	L <sub>1</sub>	F124 L <sub>3</sub> C <sub>1</sub> S <sub>1</sub>
D17	N <sub>1</sub> A <sub>1</sub> H <sub>1</sub>	P <sub>1</sub> D <sub>1</sub> A <sub>3</sub> N <sub>3</sub> S <sub>1</sub>	W71 R <sub>1</sub> G <sub>1</sub>	K <sub>1</sub>	R125
G18	C <sub>2</sub> A <sub>1</sub>	K <sub>1</sub> G <sub>1</sub>	R72 K <sub>1</sub> G <sub>1</sub>	E <sub>1</sub>	K126 R <sub>1</sub> T <sub>1</sub>
P19	T <sub>1</sub> Q <sub>1</sub>		K73 Q <sub>3</sub> L <sub>2</sub> M <sub>2</sub> N <sub>1</sub>	F <sub>2</sub>	Y127 F <sub>1</sub>
E <sub>2</sub> R <sub>36</sub>	<b>K20</b> R <sub>358</sub> Q <sub>2</sub> E <sub>1</sub>		<b>L74</b> V <sub>184</sub> I <sub>52</sub>		T128 S <sub>5</sub> A <sub>1</sub>
G <sub>1</sub> F <sub>1</sub> I <sub>6</sub>	V21 I <sub>33</sub> G <sub>1</sub>	L <sub>1</sub>	<b>Y75</b> M <sub>41</sub> I <sub>29</sub> T <sub>23</sub> A <sub>11</sub> S <sub>1</sub> L <sub>4</sub>		A129 V <sub>1</sub> G <sub>1</sub>
R <sub>2</sub> N <sub>6</sub>	K22 N <sub>4</sub> R <sub>2</sub> I <sub>1</sub>		D76 N <sub>1</sub>		F130 L <sub>2</sub>
P <sub>1</sub> H <sub>1</sub>	Q23 H <sub>2</sub> R <sub>1</sub> P <sub>1</sub>		<b>F77</b> L <sub>28</sub> S <sub>1</sub> Y <sub>1</sub>		T131 A <sub>1</sub>
W24	K <sub>1</sub>		R78 T <sub>1</sub>		I132 V <sub>7</sub> M <sub>1</sub> L <sub>1</sub> Q <sub>1</sub>
P25	A <sub>2</sub> T <sub>1</sub>		E79 D <sub>3</sub>	S <sub>1</sub>	P133 L <sub>2</sub>
F <sub>2</sub>	L26 F <sub>3</sub> S <sub>1</sub>	F <sub>1</sub>	L80	R <sub>1</sub> C <sub>2</sub>	S134 T <sub>1</sub> N <sub>1</sub> W <sub>1</sub> G <sub>1</sub>
Q <sub>1</sub> I <sub>1</sub> P <sub>4</sub> A <sub>6</sub> S <sub>6</sub>	<b>T27</b> S <sub>29</sub> P <sub>4</sub> A <sub>2</sub>	Y <sub>1</sub>	N81 K <sub>1</sub>	P <sub>1</sub> Q <sub>2</sub> K <sub>3</sub> M <sub>3</sub> L <sub>2</sub> R <sub>14</sub> V <sub>27</sub> T <sub>143</sub>	<b>I135</b> T <sub>820</sub> V <sub>141</sub> L <sub>74</sub> M <sub>33</sub> R <sub>28</sub> K <sub>6</sub> A <sub>3</sub>
Q <sub>1</sub> K <sub>1</sub> A <sub>1</sub>	<b>E28</b> A <sub>38</sub> K <sub>27</sub> G <sub>2</sub> D <sub>1</sub>	R <sub>3</sub>	K82 R <sub>8</sub> N <sub>1</sub> M <sub>1</sub>		N136 T <sub>11</sub> I <sub>2</sub>
G <sub>1</sub> D <sub>1</sub>	E29 V <sub>1</sub> G <sub>1</sub>	N <sub>1</sub> K <sub>121</sub>	<b>R83</b> K <sub>331</sub> T <sub>1</sub> I <sub>1</sub>		N137 S <sub>8</sub> H <sub>1</sub>
K30	Q <sub>3</sub> R <sub>2</sub> E <sub>1</sub> T <sub>1</sub>		T84 P <sub>2</sub> A <sub>1</sub> S <sub>1</sub>	R <sub>1</sub> K <sub>2</sub> G <sub>2</sub> A <sub>16</sub>	<b>E138</b> A <sub>89</sub> Q <sub>16</sub> G <sub>13</sub> K <sub>3</sub> T <sub>3</sub> S <sub>1</sub> D <sub>1</sub> R <sub>1</sub>
I31	L <sub>11</sub> M <sub>1</sub> V <sub>1</sub>		Q85	V <sub>1</sub> M <sub>1</sub> A <sub>1</sub> R <sub>2</sub> I <sub>2</sub> P <sub>3</sub>	<b>T139</b> A <sub>15</sub> R <sub>14</sub> K <sub>10</sub> M <sub>7</sub> Q <sub>6</sub> S <sub>2</sub> V <sub>3</sub> I <sub>2</sub> P <sub>3</sub>
N <sub>1</sub> E <sub>1</sub> R <sub>6</sub>	K32 R <sub>16</sub> E <sub>15</sub> N <sub>4</sub> I <sub>6</sub> Q <sub>3</sub> T <sub>2</sub> L <sub>1</sub>	H <sub>1</sub>	D86 E <sub>1</sub>	L <sub>1</sub>	P140 Q <sub>2</sub> T <sub>2</sub>
A33	G <sub>2</sub> V <sub>1</sub> T <sub>1</sub>	L <sub>5</sub>	F87 L <sub>3</sub>		<b>G141</b> R <sub>2</sub> A <sub>1</sub> E <sub>1</sub>
F <sub>1</sub>	L34 I <sub>6</sub> F <sub>4</sub> Y <sub>1</sub>	C <sub>5</sub>	W88 C <sub>14</sub> S <sub>7</sub>	S <sub>1</sub> N <sub>1</sub> F <sub>1</sub> A <sub>1</sub> K <sub>3</sub> L <sub>3</sub> M <sub>11</sub> V <sub>14</sub> T <sub>25</sub>	I142 V <sub>152</sub> T <sub>34</sub> M <sub>18</sub> K <sub>11</sub> L <sub>5</sub> Q <sub>3</sub> R <sub>2</sub> S <sub>1</sub> A <sub>1</sub> F <sub>1</sub>
K <sub>1</sub> F <sub>1</sub> E <sub>1</sub> A <sub>1</sub> L <sub>7</sub> M <sub>14</sub> T <sub>43</sub> I <sub>95</sub>	V35 I <sub>233</sub> T <sub>201</sub> M <sub>142</sub> L <sub>43</sub> E <sub>17</sub> R <sub>14</sub> A <sub>3</sub> K <sub>2</sub> P <sub>4</sub> S <sub>1</sub> F <sub>1</sub>	A <sub>1</sub> I <sub>6</sub>	<b>E89</b> G <sub>1</sub> D <sub>1</sub>	S <sub>1</sub>	R143 T <sub>1</sub> S <sub>1</sub> K <sub>1</sub> G <sub>1</sub>
K <sub>1</sub> G <sub>1</sub> A <sub>3</sub> D <sub>2</sub>	E36 A <sub>15</sub> D <sub>4</sub> Q <sub>2</sub> K <sub>2</sub> V <sub>2</sub> Y <sub>1</sub> T <sub>1</sub> G <sub>1</sub>		<b>V90</b> I <sub>57</sub> G <sub>2</sub>	F <sub>1</sub>	Y144 F <sub>4</sub> C <sub>1</sub>
T <sub>1</sub>	I37 T <sub>1</sub> V <sub>1</sub> F <sub>1</sub>	F <sub>1</sub>	Q91 P <sub>1</sub> H <sub>1</sub>	V <sub>1</sub> K <sub>1</sub> H <sub>1</sub> E <sub>1</sub> C <sub>1</sub>	Q145 C <sub>2</sub> E <sub>1</sub> H <sub>1</sub>
L <sub>1</sub> F <sub>1</sub>	C38 W <sub>1</sub> G <sub>1</sub> V <sub>1</sub>		L92 F <sub>1</sub>	D <sub>1</sub> C <sub>1</sub> S <sub>7</sub>	<b>Y146</b> S <sub>6</sub> F <sub>6</sub> C <sub>2</sub>
Q <sub>1</sub> K <sub>1</sub> S <sub>1</sub> M <sub>1</sub> E <sub>2</sub> G <sub>3</sub> P <sub>3</sub> L <sub>6</sub> A <sub>34</sub>	<b>T39</b> A <sub>215</sub> E <sub>12</sub> K <sub>21</sub> S <sub>12</sub> L <sub>2</sub> R <sub>2</sub> D <sub>1</sub> M <sub>3</sub> N <sub>3</sub> Q <sub>2</sub> P <sub>1</sub>		G93		N147 K <sub>1</sub> I <sub>1</sub> A <sub>1</sub>
Q <sub>2</sub> N <sub>1</sub> A <sub>1</sub> K <sub>1</sub> D <sub>12</sub>	E40 D <sub>28</sub> F <sub>17</sub> K <sub>5</sub> V <sub>2</sub> L <sub>1</sub> Q <sub>2</sub> G <sub>1</sub> Y <sub>1</sub>		I94 L <sub>5</sub>		V148 G <sub>4</sub> A <sub>3</sub>
W <sub>1</sub> L <sub>5</sub>	<b>M41</b> L <sub>747</sub> I <sub>2</sub> F <sub>1</sub>		P95		L149 P <sub>1</sub> F <sub>1</sub>
K <sub>2</sub>	E42 Q <sub>1</sub> G <sub>1</sub>		H96 Y <sub>3</sub>	H <sub>1</sub> S <sub>3</sub>	P150 S <sub>8</sub> L <sub>2</sub> A <sub>1</sub> H <sub>1</sub>
Q <sub>1</sub> E <sub>1</sub> R <sub>3</sub>	<b>K43</b> E <sub>133</sub> Q <sub>24</sub> N <sub>32</sub> R <sub>7</sub> A <sub>4</sub> T <sub>2</sub> S <sub>3</sub> H <sub>2</sub> D <sub>2</sub>	S <sub>1</sub> L <sub>1</sub>	P97 S <sub>4</sub> L <sub>2</sub> T <sub>1</sub>		<b>Q151</b> M <sub>34</sub> L <sub>18</sub> K <sub>5</sub> H <sub>1</sub>
D <sub>1</sub> K <sub>2</sub>	<b>E44</b> D <sub>189</sub> A <sub>29</sub> G <sub>2</sub> K <sub>2</sub>	V <sub>1</sub> Q <sub>1</sub> G <sub>4</sub> S <sub>31</sub>	<b>A98</b> S <sub>147</sub> G <sub>34</sub>		G152 E <sub>1</sub>
E <sub>3</sub>	G45 R <sub>1</sub> E <sub>1</sub>		G99 R <sub>1</sub> W <sub>1</sub>		W153 L <sub>1</sub>
Q <sub>1</sub>	K46 Q <sub>4</sub> N <sub>2</sub> R <sub>1</sub>		<b>L100</b> I <sub>67</sub> Q <sub>1</sub> F <sub>1</sub> V <sub>1</sub>		K154 E <sub>2</sub> Q <sub>1</sub>
F <sub>1</sub>	I47 L <sub>1</sub> F <sub>1</sub>	I <sub>1</sub> E <sub>2</sub> R <sub>18</sub>	<b>K101</b> Q <sub>25</sub> E <sub>73</sub> R <sub>42</sub> P <sub>33</sub> H <sub>18</sub> N <sub>7</sub> T <sub>4</sub> I <sub>2</sub> S <sub>2</sub> G <sub>2</sub> D <sub>2</sub> A <sub>2</sub>		G155
L <sub>1</sub> E <sub>4</sub> T <sub>2</sub>	S48 T <sub>40</sub> E <sub>25</sub> P <sub>2</sub> L <sub>1</sub> Y <sub>1</sub>	M <sub>1</sub> E <sub>2</sub> R <sub>4</sub> Q <sub>29</sub>	K102 Q <sub>82</sub> R <sub>38</sub> E <sub>11</sub> N <sub>2</sub> I <sub>1</sub>		S156 L <sub>2</sub>
Q <sub>1</sub> E <sub>1</sub> R <sub>48</sub>	K49 R <sub>187</sub> E <sub>4</sub> Q <sub>2</sub> P <sub>1</sub>	N <sub>1</sub> E <sub>1</sub> R <sub>2</sub>	<b>K103</b> N <sub>477</sub> R <sub>22</sub> S <sub>6</sub> T <sub>2</sub> Q <sub>2</sub> I <sub>1</sub> E <sub>1</sub> H <sub>1</sub> M <sub>1</sub>		P157 A <sub>4</sub> S <sub>1</sub>
W <sub>1</sub> T <sub>1</sub> V <sub>23</sub>	I50 V <sub>41</sub> L <sub>1</sub> T <sub>2</sub> M <sub>1</sub>	Q <sub>1</sub> I <sub>2</sub> N <sub>3</sub> R <sub>3</sub>	K104 N <sub>20</sub> R <sub>20</sub> Q <sub>4</sub> H <sub>1</sub> T <sub>1</sub> I <sub>1</sub>	S <sub>6</sub>	A158 S <sub>25</sub> T <sub>2</sub> P <sub>2</sub>
G51	V <sub>1</sub>	A <sub>1</sub> L <sub>1</sub>	S105 A <sub>4</sub> T <sub>2</sub>	V <sub>2</sub>	I159 L <sub>7</sub> V <sub>1</sub>
P52		G <sub>1</sub> A <sub>1</sub> I <sub>6</sub>	<b>V106</b> A <sub>58</sub> I <sub>12</sub> M <sub>9</sub> L <sub>2</sub> G <sub>1</sub>		F160 S <sub>5</sub> L <sub>1</sub>
D <sub>6</sub>	E53 D <sub>34</sub> V <sub>1</sub>	P <sub>2</sub>	T107 P <sub>7</sub> A <sub>1</sub> S <sub>1</sub>	S <sub>1</sub> K <sub>1</sub> H <sub>4</sub>	Q161 R <sub>1</sub> K <sub>1</sub> S <sub>1</sub> P <sub>1</sub>
N54	Y <sub>1</sub>	I <sub>1</sub>	<b>V108</b> I <sub>84</sub> A <sub>1</sub> L <sub>1</sub>	W <sub>1</sub> V <sub>1</sub> T <sub>1</sub> I <sub>1</sub> H <sub>1</sub> D <sub>2</sub> Y <sub>4</sub> R <sub>6</sub> G <sub>7</sub> A <sub>9</sub> C <sub>47</sub>	<b>S162</b> C <sub>189</sub> D <sub>48</sub> Y <sub>42</sub> N <sub>35</sub> A <sub>27</sub> G <sub>21</sub> H <sub>31</sub> T <sub>4</sub> W <sub>3</sub> R <sub>1</sub> V <sub>1</sub>

FIG. 2. Polymorphisms and mutations in HIV-1 RT from drug-naïve and RTI-treated patients. Amino acid mutations in HIV-1 RT from 457 plasma samples from drug-naïve patients and 1,779 samples from RTI-treated patients are reported. The clade B consensus sequence is shown as a reference. Residues associated with RTI treatment by previous in vitro or in vivo studies (24, 35, 55, 56) are underlined. Residues whose wild-type prevalence differs significantly ( $P < 0.05$ ) between naïve and treated patients are shown in boldface. The predominant mutation for each position is given closest to the reference. Mutations are represented by one-letter amino acid symbols and are numbered.

Naive	Treated	Naive	Treated	Naive	Treated
$T_1 R_1 G_1 A_2$	<b>S163</b> $G_2 I_1 R_1$	<b>P217</b> $Q_6$		$F_2$	<b>Y271</b> $S_1 F_1 H_1 D_1$
	<b>M164</b> $L_1 R_1$	$A_1$	<b>D218</b> $E_{139} T_4 H_1 Y_1 I_1 V_1$	$T_1 G_1 S_1 P_{338}$	<b>A272</b> $P_{1407} S_7 T_3 G_1$
$K_2 I_{14}$	<b>T165</b> $I_{41} L_2 A_2 P_2 V_1 M_1$	$N_1 Q_4$	<b>K219</b> $Q_{265} E_{710} R_{310} N_{310} W_5 H_3 G_2 D_2 Y_1 V_1 T_1 I_1$		<b>G273</b>
$T_1 Q_3 R_{59}$	<b>K166</b> $R_{241} Q_2 T_1 I_3 S_1 G_1 E_1$	$N_2$	<b>K220</b> $N_2 R_1 E_1$	$V_1 L_2$	<b>I274</b> $V_2 L_4$
	<b>I167</b> $V_2 F_2 N_2 M_2$		<b>H221</b> $Y_{81} L_2 Q_2 F_1 K_1 I_1$	$E_2 Q_2 R_{11}$	<b>K275</b> $R_{33} Q_{10} T_1 N_1 E_1$
	<b>L168</b>		<b>Q222</b> $H_1 R_1 L_1$	$A_1 T_2 I_{23}$	<b>V276</b> $I_{31} T_3 A_1 G_1 L_1$
$Q_1 K_2 D_{32}$	<b>E169</b> $D_{112} A_{110} K_8 N_1 G_1$		<b>K223</b> $Q_{40} E_{23} T_4 H_1$	$T_1 Q_2 R_{146}$	<b>K277</b> $R_{403} Q_2 S_1 P_1 T_1 M_1$
	<b>P170</b> $S_4 L_2 H_1$	$D_1$	<b>E224</b> $D_{12} K_1 G_1$	$R_1 P_1 N_1 H_1 E_1$	<b>Q278</b> $H_{54} E_{12} N_8 R_3 A_2 P_1 L_1$
$Y_1 L_2$	<b>F171</b> $Y_{10} L_1 I_1 C_1$		<b>P225</b> $H_{35} S_2$	$S_1$	<b>L279</b> $M_2$
	<b>R172</b> $K_{24} G_1 I_1$		<b>P226</b> $S_2$		<b>C280</b> $S_1$
$P_1 S_2 I_2 R_2 N_2 L_2 E_2 A_2 Q_2 T_{16}$	<b>K173</b> $A_{62} T_{40} E_{22} R_{310} Q_{15} N_{16} S_7 I_3 L_4 D_1$		<b>F227</b> $L_{26} Y_2$	$T_1 N_1 I_1 R_{131}$	<b>K281</b> $R_{434} H_2 N_1 P_1$
$T_1 E_1 G_2 H_2 R_2 K_{20}$	<b>Q174</b> $K_{129} R_{310} H_{11} N_4 E_2 L_1$	$H_1$	<b>L228</b> $H_{151} R_{43} N_4 Q_4 F_3 I_3 V_2 K_1$	$R_1$	<b>L282</b> $C_1 I_1$
$Y_2 H_3$	<b>N175</b> $Y_{13} H_4 I_2 T_1$		<b>W229</b>	$P_1 I_{25}$	<b>L283</b> $I_{115} F_7 P_1$
$T_1 S_2 Q_4$	<b>P176</b> $Q_2 L_2 S_2$		<b>M230</b> $L_{12}$	$K_4$	<b>R284</b> $K_{131} M_4 L_4$
$G_4 N_7 E_{85}$	<b>D177</b> $E_{374} N_{24} G_{30} S_2 H_1 Q_1 K_1 A_1$		<b>G231</b> $K_1$		<b>G285</b> $E_2 V_1 A_1 R_1$
$N_1 V_{14} L_{27} M_{46}$	<b>I178</b> $M_{143} L_{134} V_{17} F_4$		<b>Y232</b> $H_2$	$P_{12} A_{127}$	<b>T286</b> $A_{427} P_{422} V_8 N_3 S_2 G_1 L_4$
$T_1 G_1 E_1 D_4 I_{14}$	<b>Y179</b> $I_{149} D_{14} E_2 G_2 F_4 A_2 T_2$		<b>E233</b> $R_1 K_1 G_1$	$R_1$	<b>K287</b> $R_2 Q_2$
$V_6$	<b>I180</b> $V_{13} L_4 N_1 F_1$		<b>L234</b>	$Q_1 G_2 T_4 S_{15}$	<b>A288</b> $S_{86} T_{46} G_2 P_4$
	<b>Y181</b> $C_{281} I_{118} F_7 V_4 S_2 D_2 N_2 H_1$		<b>H235</b> $Y_1$	$F_2$	<b>L289</b> $I_3 P_1$
	<b>Q182</b> $H_1$		<b>P236</b> $H_1 S_1$	$S_1 P_1 A_1$	<b>T290</b> $A_1 R_1$
$C_1$	<b>Y183</b> $S_1$	$N_1 E_1$	<b>D237</b> $N_6 E_2 Y_1$	$K_1 D_{29}$	<b>E291</b> $D_{156} A_3 R_1 K_1$
$V_4$	<b>M184</b> $V_{1027} I_{12} C_2$		<b>K238</b> $T_{51} N_{10} R_2$	$W_1 T_1 A_1 I_{40}$	<b>V292</b> $I_{219} M_1 T_1 R_1$
	<b>D185</b> $H_1$		<b>W239</b> $G_1 R_1 C_1$	$L_2 V_{139}$	<b>I293</b> $V_{375} L_2 M_2 T_1$
	<b>D186</b> $E_1 Q_1$		<b>T240</b> $P_1 A_1$	$H_1 Q_4 A_{14} S_{20} T_{54}$	<b>P294</b> $T_{147} A_{102} S_{81} Q_{27} E_2 V_3 H_1 L_4 G_1$
	<b>L187</b> $M_1 P_1 S_1$		<b>V241</b> $I_6 G_2 L_4$	$M_1 K_1$	<b>L295</b> $M_2 F_2 Q_1$
$S_1$	<b>Y188</b> $L_{50} F_7 H_4 N_1 C_1$		<b>Q242</b> $H_8 R_4 L_2 K_1$	$A_1 S_{15}$	<b>T296</b> $S_{47} P_2 A_2$
$I_1$	<b>Y189</b> $I_{28}$	$S_1$	<b>P243</b> $S_7 T_2 Q_2$	$P_1 D_1 I_1 V_4 T_{11} Q_{15} A_{38} R_{46} K_{180}$	<b>E297</b> $K_{253} R_{272} A_{158} T_{59} V_{12} P_8 L_2 H_3 N_3 G_3 D_3 M_1 S_1 I_1$
	<b>G190</b> $A_{208} S_{40} R_2 Q_2 E_2 T_1 V_1$	$V_5$	<b>I244</b> $V_6 M_1$		<b>E298</b> $D_2 G_2 A_1 K_1$
	<b>S191</b>	$R_1 D_1 L_2 T_4 I_4 K_{118} Q_{23} M_{30} E_{62}$	<b>V245</b> $E_{240} Q_{114} K_{104} M_{75} T_{33} I_{14} L_4 R_2 A_2 S_2 D_1 H_1$		<b>A299</b>
$E_3$	<b>D192</b> $E_1 N_1$	$M_1$	<b>L246</b> $P_1$	$Q_1$	<b>E300</b> $Q_4 K_2 D_2 A_1$
	<b>L193</b> $F_2$	$L_1$	<b>P247</b>	$T_1 Q_1 M_1 I_{12}$	<b>L301</b> $I_{27} M_4 F_1 S_1$
	<b>E194</b> $D_2 K_2 Q_4 G_2 A_1 N_1$	$V_1 Q_2 N_2 D_{23}$	<b>E248</b> $D_{120} N_{11} K_4 A_1 Q_1 G_1$	$G_1 D_1$	<b>E302</b> $V_2 D_2 G_1$
$T_1 L_7$	<b>I195</b> $L_{11} M_1 T_1 K_1$	$R_1 N_1 L_2 Q_7$	<b>K249</b> $Q_{33} L_2 R_2 I_1$	$F_1$	<b>L303</b> $Q_3 M_2 R_2 P_1$
$R_1 D_2 E_{74}$	<b>G196</b> $E_{411} K_{11} R_4 D_3$	$N_1 G_1 E_{18}$	<b>D250</b> $E_{27} K_2 A_1 G_1$	$P_1 E_2$	<b>A304</b> $E_{12} G_1 S_1 Q_1$
$R_1 H_1 L_2 E_2 P_4 K_8$	<b>Q197</b> $K_{48} E_{31} P_{17} L_2 H_2 R_4$	$H_1 D_1 I_2 N_4$	<b>S251</b> $N_{11} L_4 H_4 Q_1 D_2 S_2 G_1 C_1$	$D_1$	<b>E305</b> $K_2 Q_2 G_2 A_1 P_1 V_1$
$Q_1 R_1$	<b>H198</b> $D_1 Q_1$		<b>W252</b> $R_1$		<b>N306</b> $D_3 K_1 S_1 I_1$
	<b>R199</b> $G_2 I_1$	$S_1 N_1$	<b>T253</b> $N_2 S_2 T_1 C_1$	$K_1$	<b>R307</b> $K_6 S_2 T_2 W_1$
$R_1 M_1 K_1 P_1 V_4 E_4 I_{31} A_{78}$	<b>T200</b> $A_{374} I_{71} E_{24} K_{10} V_5 M_2 S_2 L_2 P_2 R_2 N_1$		<b>V254</b> $I_2 A_2$	$G_1$	<b>E308</b> $D_2 Q_2 A_1$
$R_1$	<b>K201</b> $T_1 R_2$		<b>N255</b> $H_1$	$V_1 N_1 L_2$	<b>I309</b> $F_1$
$M_1 L_1 V_{46}$	<b>I202</b> $V_{186} M_2 T_2 L_1$		<b>D256</b> $G_1$	$I_1$	<b>L310</b> $I_{18} V_1 Q_1$
$G_1$	<b>E203</b> $K_{81} D_{41} A_{11} V_2 Q_4 N_2 G_1 T_1$	$V_1 L_2$	<b>I257</b> $L_4 V_1$	$T_1 R_{48}$	<b>K311</b> $R_{145} N_2 E_1 I_1$
$A_1 Q_2 D_3$	<b>E204</b> $K_2 D_2 Q_2 N_2 T_2 A_1$	$H_1$	<b>Q258</b>	$V_1 K_1 A_1 G_4 D_2$	<b>E312</b> $Q_{21} D_2 G_2 A_4 K_1 R_1 H_1$
	<b>L205</b> $F_2$	$N_1$	<b>K259</b> $E_1$	$Q_2 T_4$	<b>P313</b> $S_{17} T_{10} Q_2 P_2 A_1 V_1 E_1$
$T_1$	<b>R206</b> $G_1 T_1$	$I_1$	<b>L260</b> $I_1 V_1$	$L_1 I_2 A_2$	<b>V314</b> $L_2 I_2 A_2 E_1$
$i_1 N_1 L_1 R_2 G_4 D_4 K_8 H_7 A_2 E_{53}$	<b>Q207</b> $E_{307} A_{34} K_{28} D_{25} H_{22} R_{12} N_6 G_2 T_2 P_2 V_1$	$I_3$	<b>V261</b> $L_4 L_4$		<b>H315</b> $Q_4 M_2 L_2 W_1$
$Y_1 L_1$	<b>H208</b> $Y_{157} F_2 Q_1$	$E_2$	<b>G262</b>		<b>G316</b> $W_1 E_1$
	<b>L209</b> $W_1 Q_1$		<b>K263</b> $E_1$	$W_1 G_2 C_2 A_{21}$	<b>V317</b> $A_{121} W_3 M_1 T_2 G_2 D_1 L_4 S_1 C_1$
$W_2 V_2 M_2 F_3$	<b>L210</b> $W_{400} F_{10} S_6 V_3 M_2 R_2 T_2 G_1$	$F_1$	<b>L264</b> $F_2 W_1$	$S_1 C_2 F_2 D_3$	<b>Y318</b> $F_2 N_2 S_2 C_1 V_1$
$v_2 M_2 A_2 E_2 Q_2 T_2 S_2 G_{11} K_{203}$	<b>R211</b> $K_{911} Q_{21} S_{18} G_{17} T_{15} E_{14} A_{13} N_7 D_4 L_2 Y_1 M_1$		<b>N265</b> $I_1$	$T_2 S_1 P_1 I_1$	<b>Y319</b> $I_2 M_1 C_1 F_1 N_1$
	<b>W212</b> $L_2$		<b>W266</b> $C_1 G_1$	$A_2$	<b>D320</b> $A_4 E_2 N_2 G_2 T_2 I_1$
	<b>G213</b>		<b>A267</b>		
$S_1 L_{24}$	<b>E214</b> $L_{124} M_1 V_2 I_2$		<b>S268</b> $G_1 R_1$		
$V_1 N_1 K_1 C_2 A_2 S_2 E_2 Y_4 D_8$	<b>T215</b> $Y_{710} F_{187} S_{10} I_{37} V_{21} N_{12} C_{12} D_{10} A_3 H_2 G_1 R_1$		<b>Q269</b> $H_2$		
	<b>T216</b> $H_4 N_1$		<b>I270</b> $L_2 V_2$		

FIG. 2—Continued.



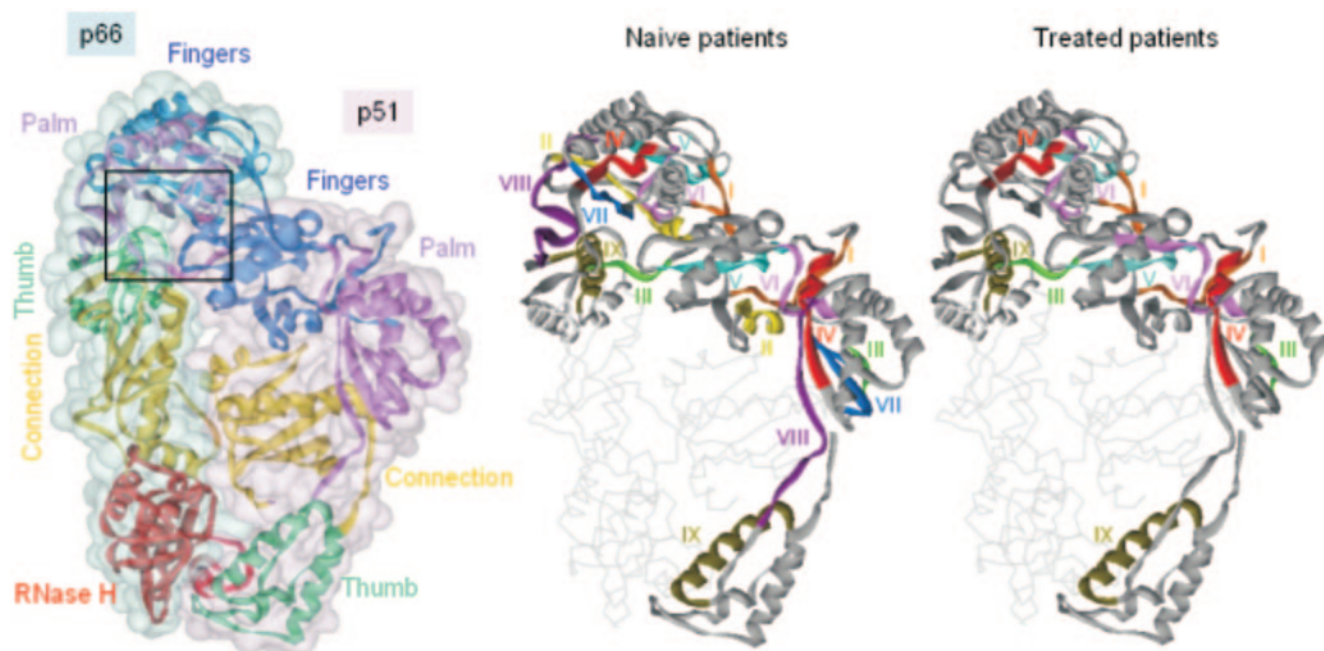


FIG. 3. HIV-1 RT conserved area locations. Left: Ribbon representation of the C $\alpha$  trace of HIV-1 RT (as found in the X-ray crystal structure with PDB code 1DLO) with p66 and p51 subunits each enveloped by a solvent-accessible surface colored in blue and pink, respectively. The boxed area is enlarged in Fig. 4. Right: Ribbon representation of the X-ray crystal structure of HIV-1 RT showing the location in both p66 and p51 subunits of the large conserved areas (I to IX) characterized in both drug-naïve and RTI-treated patients.

cantly increased in treated patients ( $P < 0.05$  to  $< 0.0001$ ). Surprisingly, a few positions were characterized by a decreased rate of mutations in treated versus drug-naïve patients: T27P (from 1.3 to 0.22%,  $P = 0.007$ ), T27A (from 1.3 to 0.11%,  $P = 0.001$ ), V35I (from 20.8 to 11.3%,  $P < 0.0001$ ), I50V (from 5 to 2.3%,  $P = 0.004$ ), R83K (from 33 to 18.6%,  $P < 0.0001$ ), and V276I (from 5 to 1.7%,  $P < 0.0001$ ).

Interestingly, among mutations not yet associated with resistance to any known RTI, some of them (T39A, K122E, and T286A) were also included as candidate mutations for drug resistance and positive reproductive fitness, by having a ratio of nonsynonymous mutations to synonymous mutations ( $K_a/K_s$ ) of  $> 1$ , with values similar to those observed for the already-known RTI resistance mutations (12).

**Structural interpretation of the minimal conserved areas of HIV-1 RT.** It is known that the HIV-1 RT catalytic activities are localized in the p66 subunit that is composed of all 560 aa encoded in the RT gene, whereas the p51 subunit, made up of 440 shared aa, displays no enzymatic activity although it provides essential structural support to the enzymatically active p66 monomer in the RT dimer (25). It is important to recall that the shared 440 aa are arranged in structurally different ways in p66 and p51 and also that mutations occurring in vivo in the RT genome influence both subunits (50). Therefore, in an attempt to define an improved minimal conserved structure of HIV-1 RT, we rationalize our findings also in structural terms by mapping the first 320 aa of RT from both naïve and treated patients onto a three-dimensional representation of both p66 and p51 subunits (Fig. 3). Among the large conserved areas observed in drug-naïve patients, regions I (7 to 19), II (70 to 82), and V (124 to 134) are located in the “fingers” subdo-

main (1 to 88 and 121 to 146). Regions III (91 to 97), IV (107 to 117), VI (147 to 157), VII (181 to 194), and VIII (216 to 244) are located in the “palm” subdomain (89 to 120 and 147 to 242), while region IX (252 to 271), containing a long  $\alpha$ -helix, usually termed  $\alpha$ H, is in the “thumb” subdomain (243 to 311) (Fig. 3). The IV and the VII areas contain, respectively, the three aspartate residues involved in catalysis (D110, D185, and D186). The largest conserved area in drug-naïve patients, VIII (216 to 244), contains residues (224 to 235) that are involved in positioning the primer terminus (“primer grip” [40]).

Several known mutations associated with RTI treatment and/or RTI resistance (at residues K219, H221, K223, P225, F227, L228, and K238) appeared in the VIII area, whereas the  $\beta$ -stranded stretch 229 to 237 and residues 224, 226, and 239 to 244 remained fully conserved (Fig. 4). The fact that in this region some residues that also line the NNRTI binding pocket (W229, E233, L234, and P236) are not mutated suggests that the primer grip residues have critical architectural and functional roles.

A “template grip” has also been described in the fingers subdomain (residues 67 to 83) and in the palm (residues 86 to 93 and 148 to 154) that interacts with the DNA sugar-phosphate backbone of the template strand (9, 32, 33, 60). In naïve patients, with the exception of D67, S68, T69, and V90 (variability of  $< 5\%$ ) and R83 (variability of 33%), all these residues are conserved and some are clustered in conserved regions (II and VI). In drug-treated patients, many NRTI resistance mutations fall into these areas (K65R, D67N, T69D, K70R, L74V, V75I, F77L, and Q151M), whereas residues 66, 71 to 73, 76, the  $\alpha$ B 78 to 82, 86 and 87, 89, 91 to 93, 147 to 150, and 152 to 154 remain conserved. It is interesting that, from mutational

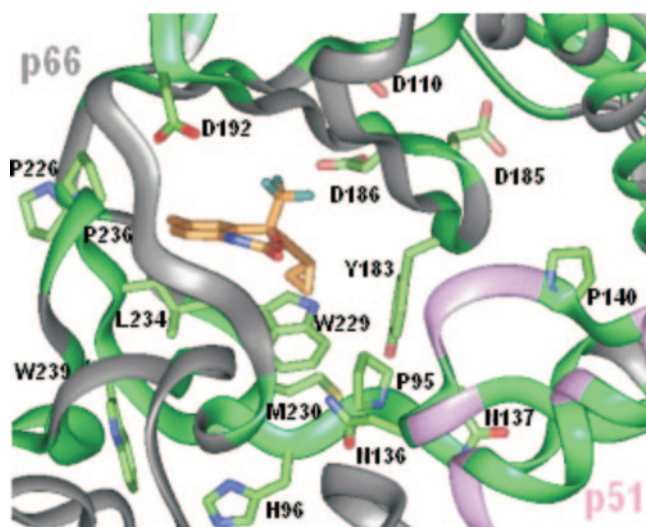


FIG. 4. Detail of the NNRTI binding and catalytic sites in HIV-1 RT. Enlarged view of the boxed area shown in Fig. 3. The  $\alpha$  trace of HIV-1 RT, as found in its crystallographic complex (PDB code 1FK9) with efavirenz (shown as sticks), is displayed as a ribbon, colored in gray and pink for the p66 and p51 subunits, respectively, except for all residues having a mutation frequency of  $<1\%$ , which are colored in green. Side chains of nonmutable residues in the neighborhood of the NNRTI binding site, together with the active-site aspartates (D185, D186, and D110), are shown as sticks, with carbon atoms colored in green.

analyses of the 67 to 81 region, residues W71, R72, D76, R78, and N81 (residues also evolutionarily conserved) were intolerant of substitution compared to the other nine, thus suggesting an important role for these residues in polymerase activity (8, 37).

Region IX (252 to 271) is the largest of all the fully conserved stretches in both naïve and drug-treated patients (Fig. 1 and 3). Residues 259 to 284 exhibit sequence homology with other nucleic acid polymerases and have been termed the “helix clamp” (4, 28). In p66 these residues belong to the polymerase domain as part of the thumb, and it has been shown by alanine-scanning mutagenesis that specific amino acids (Q258, G262, and W266) are critical for template-primer affinity (5). Another recent paper has demonstrated that mutation of N255 or N265, each adjoining the minor groove-binding track, to aspartate leads, respectively, to a drastic reduction of processive synthesis on both templates or to a loss of processive polymerization on DNA but not on RNA (18). The invariant region IX in the p66 subunit is essential for DNA polymerization and can have an important role also in p51. In fact,  $\alpha$ H (V254 to S268) in the p51 subunit is in contact with the RNase H domain of p66, suggesting a substantial contribution of this region to dimer stabilization (38, 41, 62).

In this regard, the segment from I135 to P140 in p51, which is part of the  $\beta$ 7- $\beta$ 8 loop and essential for the catalytic activity of the p66 subunit, has been identified as a “hot spot” of binding energy (53) and is invariant in HIV-1 RT in three residues, N136, N137, and P140. In particular, N136 represents a highly conserved amino acid among all known lentivirus RTs, whose role in HIV-1 RT has been recently shown to be severely compromised upon mutation to other amino acids (2a).

The NNRTI binds at a hydrophobic pocket approximately 10 Å from the polymerase active site (38). Analyzing the conservation within the NNRTI binding pocket, in drug-naïve patients, 14 of 19 residues that make up the hydrophobic NNRTI binding pocket (P95, L100, K103, V108, Y181, Y183, Y188, G190, P225, F227, W229, E233, L234, and P236) are fully conserved ( $<1\%$ ), others (K101, V106, and Y318) are mutated at low frequency ( $<3\%$ ), and two (A98 and V179) are mutated in 5 to 8% of patients (Fig. 4). As expected, and in contrast to what is observed in naïve patients, the NNRTI binding cavity of HIV-1 RT from drug-experienced individuals displays a high and variable occurrence (from 5 to 29%) of mutations at positions A98, K101, K103, V106, V108, V179, Y181, and G190. Nevertheless, seven residues (P95, Y183, W229, E233, L234, P236, and Y318) remain fully conserved, and others (L100, Y188, P225, and F227) are mutated only rarely ( $<4\%$  frequency) (Fig. 4). The invariant amino acids in the HIV-1 NNRTI binding pocket, as already shown for W229 and Y318 (44, 45), should be regarded as prime candidates to be targeted by modified or novel NNRTIs (43) in order to rationally design new drugs with a more favorable resistance profile.

**RT evolutionary conservation.** Being aware of the potential limitation of working solely with HIV-1-derived data (mainly B subtype), we undertook an analysis of RT conservation also in a variety of HIV-1 clades and other published retroviral sequences. Consistent with the pattern of sequence conservation observed in HIV-1 RT from isolates of drug-naïve patients, the largest invariant regions are also highly conserved in RTs from other HIV-1 subtypes (A, C, D, F1, F2, G, H, J, and K), with the single exception of RT from clade H in which, at position 107, Thr (present in all subtypes) has been replaced by Ser (a conservative mutation) (data not shown). This indicates that, in the absence of any drug pressure and independent of both virus origin and circulation, HIV-1 RT amino acid variations (i.e., mutations) are indeed allowed (more than 30% of the sequence), but only in restricted selected areas.

In contrast, when we compared the sequence of the first 320 aa of HIV-1 RT consensus clade B with the corresponding RT sequences from other lentiviruses (HIV-1 subtypes, HIV-2, EIAV, visna/maedi virus, CAEV, BLV, FIV, and SIV) a completely different picture (Fig. 5), characterized by a substantial decrease of amino acid conservation, became visible. Indeed, among all viruses only 73/320 residues were fully conserved, 47 residues mutated toward strongly conservative amino acid substitutions, and 17 residues mutated toward weakly conservative substitutions, with an overall evolutionary conservation of 137 residues out of 320 (43%). As previously proposed (3), the N terminus of the HIV-1 RT lies on a surface distantly located from the active site and is accessible to the HIV-1 PR. This area is subject to a considerable variation. Instead, the most evolutionarily conserved RT regions (in the frame of aa 20 to 190 of HIV-1 RT) are within the inner surface of the enzyme and responsible for polymerization function. Overall, several evolutionarily conserved stretches in lentiviral RTs were also found invariant in HIV-1 RT under antiviral pressure (G18-P19, Q23 to L26, E29-K30, G45 to I47, N57 to P59, W71-R72, R78 to K82, G93 to P97, L109 to V111, D113 to Y115, P119-L120, Y127 to I132, N136-N137, R143-Y144, L149-P150, G152-W153, S156-P157, I159 to Q161, I167-L168, D185 to



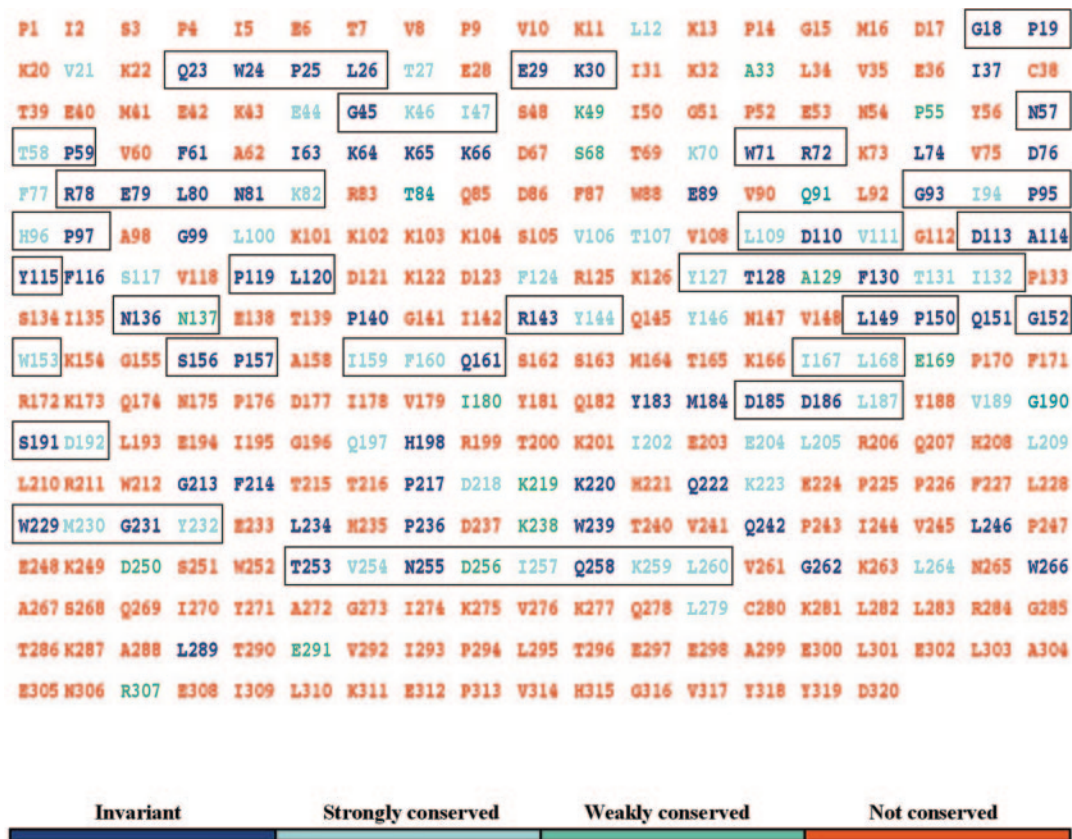


FIG. 5. Conserved regions of RT in lentiviruses. The sequence of the first 320 aa of HIV-1 RT of clade B consensus (shown as a reference) is colored according to the degree of conservation observed in different lentiviruses. Published *pol* consensus sequences of pure HIV-1 subtypes (A, B, C, D, F1, F2, G, H, J, and K) together with published *pol* sequences of HIV-2 (P18042 from residue 198), visna virus (AAM51650 from residue 160), FIV (AAM13444 from residue 163), EIAV (AAK21112 from residue 175), CAEV (AAG48629 from residue 147), ovine lentivirus (AAO33139 from residue 14), puma lentivirus (AAA67168 from residue 64), SIVS4 (P12502 from residue 168), SIVCZ (P17283 from residue 180), and BLV (P19561 from residue 157) were aligned using the Clustal X program. Mutated residues with conservative amino acid substitutions were considered either strongly conserved, if amino acid alterations were from/to one of the amino acids within the groups STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, and FYW, or weakly conserved, if alterations were from/to one of the amino acids within the groups CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEHQK, MILVF, and HFY. RT conserved regions in lentiviruses and HIV-1 (from our cohort of drug-naïve and RTI-treated patients) are boxed.

L187, S191-D192, W229 to Y232, and T253 to L260) (Fig. 5, see boxes), whereas a lower number of RT residues are overall conserved in more distantly related retroviral sequences (P52, N57, P59, K65, R72, D76, R78, N81, D110, D113, F116, A129-F130, L149 to G152, S156-P157, L168, Y183, D185-D186, L209, G213, K220, G231, and G262) (34).

To our knowledge, among the 202 HIV-1 RT residues identified as invariant in our study, 62 have not been yet individually characterized for their function and should become the subject of future therapeutically oriented studies (P1 to S3, I5, V8 to V10, L12 to P19, K22 to W24, K82, T84 to F87, E89, L205-R206, L209, W212-G213, T216-P217, K220, Q222, P236-D237, W239 to I244, L246-P247, W252, G273-I274, T290, L295, E298, E300, E302, A304 to I309, V314 to G316, and Y319-D320). Some of these residues are invariant among all other lentiviral RTs, suggesting an essential role for RT structure and/or function (G18-P19, Q23-W24, E89, G213, P217, K220, Q222, P236, W239, Q242, and L246) (Fig. 5).

At least 140 HIV-1 RT conserved residues have been mutagenized in vitro, and the effect of one or more of these

mutations has been analyzed by measuring RNA- and/or DNA-dependent DNA polymerase activity, RNase H activity, strand transfer, structural stability, and/or p66-p51 heterodimerization (2, 2a, 4, 5, 7, 8, 11, 18, 20–22, 37, 40–42, 44, 45, 59, 64). Of these 140 mutagenized residues, 79 (56%) have been shown to be extremely important to keep the functional and/or structural integrity of RT, thus confirming their key role in RT and the necessity for their conservation. In contrast, the other 61 invariant residues (44%) did not show any drastic effects on the efficiency of in vitro RT/RNase H activity ( $\geq 30\%$  of that of wild-type HIV RT). This does not necessarily prove an unsupportive role in HIV-1 RT structure and/or function but suggests that these amino acids, in principle, may be mutated without a deleterious loss of catalytic RT efficiency. The fact, for instance, that the majority of these residues (48/61) are subject to variation in the larger world of lentiviruses (3) (Fig. 5) may confirm that substitutions on these sites are tolerable in vivo, without causing a major deleterious effect on the RT enzyme. Therefore, we cannot exclude that mutations at some of these residues may occur under the pressure of novel

RT inhibitors with a different mechanism of action. Future studies may provide answers to these considerations. On the other hand, the other residues (13/61) that are invariant also among other lentiviruses may play a role or perform a function that cannot be easily detected by the usual *in vitro* assays, which do not always cover the multifunctional activities of HIV-1 RT, but are possibly amenable to targeting by new drugs. This further emphasizes the important value of both *in vivo* and *in vitro* observations to achieve results that are both noticeable and reliable and demonstrates that the need of evolutionary conservation is not sufficient to avoid the problem of RT becoming resistant to nucleoside analogues, even if they interact with the well-conserved polymerase active site.

Finally, we should consider that the conservation or mutation at the amino acid level observed *in vivo* is also dependent, besides the functional or structural constraint, on the nature of the amino acid itself, and thus the codon usage may be relevant in the development of immune escape and drug resistance (36). To detect the overall selection pressure at the amino acid level of HIV-1 RT, nucleic acid sequences of 179 randomly selected patients with clade B *pol* subtype were analyzed. The ratio between nonsynonymous and synonymous substitutions (dn/ds) was calculated as a measure of selection pressure (66). Both RT and PR genes were under strong purifying (negative) selection pressure (dn/ds < 1) in both drug-naïve and drug-treated patients, as previously reported (12, 56). Interestingly, RT nucleotide sequences showed a remarkably lower dn/ds ratio than PR, suggesting a greater conservation of RT even under drug pressure (data not shown). Recently, analyzing 40,000 samples of sequences of the entire PR and the first 381 RT codons, Chen et al. detected 69 individual PR mutations and 142 RT mutations under positive selection pressure (12). Among the 107/381 RT codons reported with a  $K_a/K_s$  value of >1, very few were invariant residues in our study (I31, S105, T107, P176, P243, I257, A304, I309, and Y319).

In conclusion, HIV-1 can become resistant to the currently available drugs because of its high replication and mutation rate (54), yet several amino acid sequences are highly conserved even under strong drug pressure. The data presented here may allow the design of new drugs that force the virus toward a mandatory escape to residues characterized by changes in the replicative capacity extremely detrimental for its replication. The resulting mutant virus may be replicating at a limited rate such that its ability to damage the immune system, as well as to further mutate to compensate for the loss of replication capacity, would be remarkably decreased. Therefore, the characterization of the highly conserved residues (involved in protein stability, dimerization ability, template-primer binding, or catalytic activity, and some still poorly understood functions) could help in the rational design of new HIV-1 inhibitors with alternative mechanisms of action and more favorable resistance profiles.

#### ACKNOWLEDGMENTS

This work was supported by grants of the Italian Ministry of Health (Project AIDS and Current and Finalized Research), the Ministry of University and Scientific Research (FIRB), the European Community (QLKT-CT-2000-00291, and Descartes prize 2001 HPAW-2002-90001), and by one unrestricted educational grant from Glaxo-Smith-Kline, Italy.

We thank Sara Giannella, Federica Forbici, Maria Concetta Bellocchi, Alessandra Cenci, Fabio Continenza, Andrea Biddittu, and Sandro Bonfigli for sequencing and data management and all the Italian Cohort of Antiretroviral Naïve patient study group participants and members.

#### REFERENCES

- Alexander, C. S., W. Dong, K. Chan, N. Jahnke, M. V. O'Shaughnessy, T. Mo, M. A. Piaseczny, J. S. Montaner, and P. R. Harrigan. 2001. HIV protease and reverse transcriptase variation and therapy outcome in antiretroviral-naïve individuals from a large North American cohort. *AIDS* 15:601–607.
- Auwerx, J., J. Van Nieuwenhove, F. Rodríguez-Barrios, S. de Castro, S. Velázquez, F. Ceccherini-Silberstein, E. De Clercq, M. J. Camarasa, C. F. Perno, F. Gago, and J. Balzarini. 2005. The N137 and P140 amino acids in the p51 and the P95 amino acid in the p66 subunit of human immunodeficiency virus type 1 reverse transcriptase are instrumental to maintain catalytic activity and to design new classes of anti-HIV-1 drugs. *FEBS Lett.* 579:2294–2300.
- Balzarini, J., J. Auwerx, F. Rodríguez-Barrios, A. Chedad, V. Farkas, F. Ceccherini-Silberstein, C. García-Aparicio, S. Velázquez, E. De Clercq, C. F. Perno, M. J. Camarasa, F. Gago. 2005. The amino acid Asn136 in HIV-1 Reverse Transcriptase (RT) maintains efficient association of both RT subunits and enables the rational design of novel RT inhibitors. *Mol. Pharmacol.* 68:49–60.
- Barber, A. M., A. Hizi, J. V. Maizel, and S. H. Hughes. 1990. HIV-1 reverse transcriptase: structure predictions for the polymerase domain. *AIDS Res. Hum. Retrovir.* 6:1061–1072.
- Beard, W. A., D. T. Minnick, C. L. Wade, R. Prasad, R. L. Won, A. Kumar, T. A. Kunkel, and S. H. Wilson. 1996. Role of the "helix clamp" in HIV-1 reverse transcriptase catalytic cycling as revealed by alanine-scanning mutagenesis. *J. Biol. Chem.* 271:12213–12220.
- Beard, W. A., S. J. Stahl, H. R. Kim, K. Bebenek, A. Kumar, M. P. Strub, S. P. Becerra, T. A. Kunkel, and S. H. Wilson. 1994. Structure/function studies of human immunodeficiency virus type 1 reverse transcriptase. Alanine scanning mutagenesis of an alpha-helix in the thumb subdomain. *J. Biol. Chem.* 269:28091–28097.
- Boden, D., A. Hurley, L. Zhang, Y. Cao, Y. Guo, E. Jones, J. Tsay, J. Ip, C. Farthing, K. Limoli, N. Parkin, and M. Markowitz. 1999. HIV-1 drug resistance in newly infected individuals. *JAMA* 282:1135–1141.
- Boyer, P. L., A. L. Ferris, P. Clark, J. Whitmer, P. Frank, C. Tantillo, E. Arnold, and S. H. Hughes. 1994. Mutational analysis of the fingers and palm subdomains of human immunodeficiency virus type-1 (HIV-1) reverse transcriptase. *J. Mol. Biol.* 243:472–483.
- Boyer, P. L., A. L. Ferris, and S. H. Hughes. 1992. Mutational analysis of the fingers domain of human immunodeficiency virus type 1 reverse transcriptase. *J. Virol.* 66:7533–7537.
- Boyer, P. L., C. Tantillo, A. Jacobo-Molina, R. G. Nanni, J. Ding, E. Arnold, and S. H. Hughes. 1994. Sensitivity of wild-type human immunodeficiency virus type 1 reverse transcriptase to dideoxynucleotides depends on template length; the sensitivity of drug-resistant mutants does not. *Proc. Natl. Acad. Sci. USA* 91:4882–4886.
- Ceccherini-Silberstein, F., F. Erba, F. Gago, A. Bertoli, F. Forbici, M. C. Bellocchi, C. Gori, R. d'Arrigo, L. Marcon, C. Balotta, A. Antinori, A. d'Arminio Monforte, and C. F. Perno. 2004. Identification of the minimal conserved structure of HIV-1 protease in the presence and absence of drug pressure. *AIDS* 18:F11–F19.
- Chao, S. F., V. L. Chan, P. Juranka, A. H. Kaplan, R. Swanstrom, and C. A. Hutchison III. 1995. Mutational sensitivity patterns define critical residues in the palm subdomain of the reverse transcriptase of human immunodeficiency virus type 1. *Nucleic Acids Res.* 23:803–810.
- Chen, L., A. Perlina, and C. J. Lee. 2004. Positive selection detection in 40,000 human immunodeficiency virus (HIV) type 1 sequences automatically identifies drug resistance and positive fitness mutations in HIV protease and reverse transcriptase. *J. Virol.* 78:3722–3732.
- Clavel, F., and A. J. Hance. 2004. HIV drug resistance. *N. Engl. J. Med.* 350:1023–1035.
- Coffin, J. M. 1995. HIV population dynamics *in vivo*: implications for genetic variation, pathogenesis, and therapy. *Science* 267:483–489.
- D'Aquila, R. T., J. M. Schapiro, F. Brun-Vezinet, B. Clotet, B. Conway, L. M. Demeter, R. M. Grant, V. A. Johnson, D. R. Kuritzkes, C. Loveday, R. W. Shafer, and D. D. Richman. 2002. Drug resistance mutations in HIV-1. *Top. HIV Med.* 10:21–25.
- De Clercq, E. 2004. Antivirals and antiviral strategies. *Nat. Rev. Microbiol.* 2:704–720.
- EuroGuidelines Group for HIV Resistance. 2001. Clinical and laboratory guidelines for the use of HIV-1 drug resistance testing as part of treatment management: recommendations for the European setting. *AIDS* 15:309–320.
- Fisher, T. S., T. Darden, and V. R. Prasad. 2003. Mutations proximal to the minor groove-binding track of human immunodeficiency virus type 1 reverse



- transcriptase differentially affect utilization of RNA versus DNA as template. *J. Virol.* 77:5837–5845.
19. Frater, A. J., A. Beardall, K. Ariyoshi, D. Churchill, S. Galpin, J. R. Clarke, J. N. Weber, and M. O. McClure. 2001. Impact of baseline polymorphisms in RT and protease on outcome of highly active antiretroviral therapy in HIV-1-infected African patients. *AIDS* 15:1493–1502.
  20. Gao, H. Q., P. L. Boyer, E. Arnold, and S. H. Hughes. 1998. Effects of mutations in the polymerase domain on the polymerase, RNase H and strand transfer activities of human immunodeficiency virus type 1 reverse transcriptase. *J. Mol. Biol.* 277:559–572.
  21. Ghosh, M., P. S. Jacques, D. W. Rodgers, M. Ottman, J. L. Darlix, and S. F. Le Grice. 1996. Alterations to the primer grip of p66 HIV-1 reverse transcriptase and their consequences for template-primer utilization. *Biochemistry* 35:8553–8562.
  22. Goel, R., W. A. Beard, A. Kumar, J. R. Casas-Finet, M. P. Strub, S. J. Stahl, M. S. Lewis, K. Bebenek, S. P. Becerra, and T. A. Kunkel. 1993. Structure/function studies of HIV-1(1) reverse transcriptase: dimerization-defective mutant L289K. *Biochemistry* 32:13012–13018.
  23. Gonzales, M. J., R. N. Machekano, and R. W. Shafer. 2001. Human immunodeficiency virus type 1 reverse-transcriptase and protease subtypes: classification, amino acid mutation patterns, and prevalence in a northern California clinic-based population. *J. Infect. Dis.* 184:998–1006.
  24. Gonzales, M. J., T. D. Wu, J. Taylor, I. Belitskaya, R. Kantor, D. Israelski, S. Chou, A. R. Zolopa, W. J. Fessel, and R. W. Shafer. 2003. Extended spectrum of HIV-1 reverse transcriptase mutations in patients receiving multiple nucleoside analog inhibitors. *AIDS* 17:791–799.
  25. Gotte, M., X. Li, and M. A. Wainberg. 1999. HIV-1 reverse transcription: a brief overview focused on structure-function relationships among molecules involved in initiation of the reaction. *Arch. Biochem. Biophys.* 365:199–210.
  26. Grossman, Z., N. Vardinon, D. Chemtob, M. L. Alkan, Z. Bentwich, M. Burke, G. Gottesman, V. Istomin, I. Levi, S. Maayan, E. Shahar, and J. M. Schapiro. 2001. Genotypic variation of HIV-1 reverse transcriptase and protease: comparative analysis of clade C and clade B. *AIDS* 15:1453–1460.
  27. Harris, D., R. Lee, H. S. Misra, P. K. Pandey, and V. N. Pandey. 1998. The p51 subunit of human immunodeficiency virus type 1 reverse transcriptase is essential in loading the p66 subunit on the template primer. *Biochemistry* 37:5903–5908.
  28. Hermann, T., T. Meier, M. Gotte, and H. Heumann. 1994. The 'helix clamp' in HIV-1 reverse transcriptase: a new nucleic acid binding motif common in nucleic acid polymerases. *Nucleic Acids Res.* 22:4625–4633.
  29. Hirsch, M. S., F. Brun-Vezinet, B. Clotet, B. Conway, D. R. Kuritzkes, R. T. D'Aquila, L. M. Demeter, S. M. Hammer, V. A. Johnson, C. Loveday, J. W. Mellors, D. M. Jacobsen, and D. D. Richman. 2003. Antiretroviral drug resistance testing in adults infected with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society-USA Panel. *Clin. Infect. Dis.* 37:113–128.
  30. Hirsch, M. S., F. Brun-Vezinet, R. T. D'Aquila, S. M. Hammer, V. A. Johnson, D. R. Kuritzkes, C. Loveday, J. W. Mellors, B. Clotet, B. Conway, L. M. Demeter, S. Vella, D. M. Jacobsen, and D. D. Richman. 2000. Antiretroviral drug resistance testing in adult HIV-1 infection: recommendations of an International AIDS Society-USA Panel. *JAMA* 283:2417–2426.
  31. Hsiou, Y., J. Ding, K. Das, A. D. Clark, Jr., S. H. Hughes, and E. Arnold. 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7 Å resolution: implications of conformational changes for polymerization and inhibition mechanisms. *Structure* 4:853–860.
  32. Huang, H., R. Chopra, G. L. Verdine, and S. C. Harrison. 1998. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* 282:1669–1675.
  33. Jacobo-Molina, A., J. Ding, R. G. Nanni, A. D. Clark, Jr., X. Lu, C. Tantillo, R. L. Williams, G. Kamer, A. L. Ferris, and P. Clark. 1993. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. USA* 90:6320–6324.
  34. Johnson, M. S., M. A. McClure, D. F. Feng, J. Gray, and R. F. Doolittle. 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc. Natl. Acad. Sci. USA* 83:7648–7652.
  35. Johnson, V. A., F. Brun-Vezinet, B. Clotet, B. Conway, R. T. D'Aquila, L. M. Demeter, D. R. Kuritzkes, D. Pillay, J. M. Schapiro, A. Telenti, and D. D. Richman. 2004. Update of the drug resistance mutations in HIV-1: 2004. *Top. HIV Med.* 12:119–124.
  36. Kijak, G. H., J. R. Currier, S. Tovanabutra, J. H. Cox, N. L. Michael, S. A. Wegner, D. L. Birx, and F. E. McCutchan. 2004. Lost in translation: implications of HIV-1 codon usage for immune escape and drug resistance. *AIDS Rev.* 6:54–60.
  37. Kim, B., T. R. Hathaway, and L. A. Loeb. 1996. Human immunodeficiency virus reverse transcriptase. Functional mutants obtained by random mutagenesis coupled with genetic selection in *Escherichia coli*. *J. Biol. Chem.* 271:4872–4878.
  38. Kohlstaedt, L. A., J. Wang, J. M. Friedman, P. A. Rice, and T. A. Steitz. 1992. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 256:1783–1790.
  39. Little, S. J., S. Holte, J. P. Routy, E. S. Daar, M. Markowitz, A. C. Collier, R. A. Koup, J. W. Mellors, E. Connick, B. Conway, M. Kilby, L. Wang, J. M. Whitcomb, N. S. Hellmann, and D. D. Richman. 2002. Antiretroviral-drug resistance among patients recently infected with HIV. *N. Engl. J. Med.* 347:385–394.
  40. Palaniappan, C., M. Wisniewski, P. S. Jacques, S. F. Le Grice, P. J. Fay, and R. A. Bambara. 1997. Mutations within the primer grip region of HIV-1 reverse transcriptase result in loss of RNase H function. *J. Biol. Chem.* 272:11157–11164.
  41. Pandey, P. K., N. Kaushik, K. Singh, B. Sharma, A. K. Upadhyay, S. Kumar, D. Harris, and V. N. Pandey. 2002. Insertion of a small peptide of six amino acids into the beta7-beta8 loop of the p51 subunit of HIV-1 reverse transcriptase perturbs the heterodimer and affects its activities. *BMC Biochem.* 3:18.
  42. Pandey, P. K., N. Kaushik, T. T. Talele, P. N. Yadav, and V. N. Pandey. 2001. The beta7-beta8 loop of the p51 subunit in the heterodimeric (p66/p51) human immunodeficiency virus type 1 reverse transcriptase is essential for the catalytic function of the p66 subunit. *Biochemistry* 40:9505–9512.
  43. Pata, J. D., W. G. Stirtan, S. W. Goldstein, and T. A. Steitz. 2004. Structure of HIV-1 reverse transcriptase bound to an inhibitor active against mutant reverse transcriptases resistant to other nonnucleoside inhibitors. *Proc. Natl. Acad. Sci. USA* 101:10548–10553.
  44. Pelemans, H., R. Esnouf, K. L. Min, M. Parniak, E. De Clercq, and J. Balzarini. 2001. Mutations at amino acid positions 63, 189, and 396 of human immunodeficiency virus type 1 reverse transcriptase (RT) partially restore the DNA polymerase activity of a Trp229Tyr mutant RT. *Virology* 287:143–150.
  45. Pelemans, H., R. M. Esnouf, H. Jonckheere, E. De Clercq, and J. Balzarini. 1998. Mutational analysis of Tyr-318 within the non-nucleoside reverse transcriptase inhibitor binding pocket of human immunodeficiency virus type I reverse transcriptase. *J. Biol. Chem.* 273:34234–34239.
  46. Perno, C. F., A. Cozzi-Lepri, C. Balotta, A. Bertoli, M. Violin, L. Monno, T. Zauli, M. Montroni, G. Ippolito, and A. d'Arminio Monforte. 2002. Low prevalence of primary mutations associated with drug resistance in antiretroviral-naïve patients at therapy initiation. *AIDS* 16:619–624.
  47. Perno, C. F., A. Cozzi-Lepri, C. Balotta, F. Forbici, M. Violin, A. Bertoli, G. Facchi, P. Pezzotti, G. Cadeo, G. Tositti, S. Pasquinucci, S. Pauluzzi, A. Scalzini, B. Salassa, A. Vincenti, A. N. Phillips, F. Dianzani, A. Appice, G. Angarano, L. Monno, G. Ippolito, M. Mononi, and A. d'Arminio Monforte. 2001. Secondary mutations in the protease region of human immunodeficiency virus and virologic failure in drug-naïve patients treated with protease inhibitor-based therapy. *J. Infect. Dis.* 184:983–991.
  48. Perrin, L., L. Kaiser, and S. Yerly. 2003. Travel and the spread of HIV-1 genetic variants. *Lancet Infect. Dis.* 3:22–27.
  49. Ren, J., J. Milton, K. L. Weaver, S. A. Short, D. I. Stuart, and D. K. Stammers. 2000. Structural basis for the resilience of efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. *Struct. Fold Des.* 8:1089–1094.
  50. Restle, T., B. Muller, and R. S. Goody. 1992. RNase H activity of HIV reverse transcriptases is confined exclusively to the dimeric forms. *FEBS Lett.* 300:97–100.
  51. Rhee, S. Y., T. Liu, J. Ravela, M. J. Gonzales, and R. W. Shafer. 2004. Distribution of human immunodeficiency virus type 1 protease and reverse transcriptase mutation patterns in 4,183 persons undergoing genotypic resistance testing. *Antimicrob. Agents Chemother.* 48:3122–3126.
  52. Ribeiro, R. M., S. Bonhoeffer, and M. A. Nowak. 1998. The frequency of resistant mutant virus before antiretroviral therapy. *AIDS* 12:461–465.
  53. Rodriguez-Barrios, F., C. Perez, E. Lobaton, S. Velazquez, C. Chamorro, A. San-Felix, M. J. Perez-Perez, M. J. Camarasa, H. Pelemans, J. Balzarini, and F. Gago. 2001. Identification of a putative binding site for [2',5'-bis-O-(tert-butylidimethylsilyl)-β-D-ribofuranosyl]-3'-spiro-5'-(4'-amino-1',2'-oxathiole-2',2'-dioxide)thymine (TSAO) derivatives at the p51-p66 interface of HIV-1 reverse transcriptase. *J. Med. Chem.* 44:1853–1865.
  54. Sarafianos, S. G., K. Das, S. H. Hughes, and E. Arnold. 2004. Taking aim at a moving target: designing drugs to inhibit drug-resistant HIV-1 reverse transcriptases. *Curr. Opin. Struct. Biol.* 14:716–730.
  55. Schinazi, R. F., B. A. Larder, and J. Mellors. 2000. Mutations in retroviral genes associated with drug resistance: 2000–2001 update. *Int. Antivir. News* 8:65–91.
  56. Seibert, S. A., C. Y. Howell, M. K. Hughes, and A. L. Hughes. 1995. Natural selection on the gag, pol, and env genes of human immunodeficiency virus 1 (HIV-1). *Mol. Biol. Evol.* 12:803–813.
  57. Shafer, R. W. 2002. Genotypic testing for human immunodeficiency virus type 1 drug resistance. *Clin. Microbiol. Rev.* 15:247–277.
  58. Svicher, V., F. Ceccherini-Silberstein, F. Erba, M. Santoro, C. Gori, M. C. Bellocchi, S. Giannella, M. P. Trotta, A. d'Arminio Monforte, A. Antinori, and C. F. Perno. 2005. Novel human immunodeficiency virus type 1 protease mutations potentially involved in resistance to protease inhibitors. *Antimicrob. Agents Chemother.* 49:2015–2025.
  59. Tachedjian, G., H. E. Aronson, and S. P. Goff. 2000. Analysis of mutations and suppressors affecting interactions between the subunits of the HIV type 1 reverse transcriptase. *Proc. Natl. Acad. Sci. USA* 97:6334–6339.



60. Tantillo, C., J. Ding, A. Jacobo-Molina, R. G. Nanni, P. L. Boyer, S. H. Hughes, R. Pauwels, K. Andries, P. A. Janssen, and E. Arnold. 1994. Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase. Implications for mechanisms of drug inhibition and resistance. *J. Mol. Biol.* **243**:369–387.
61. Verbiest, W., S. Brown, C. Cohen, M. Conant, K. Henry, S. Hunt, M. Sension, A. Stein, R. Stryker, M. Thompson, P. Schel, R. Van Den Broeck, S. Bloor, T. Alcorn, M. Van Houtte, B. Larder, and K. Hertogs. 2001. Prevalence of HIV-1 drug resistance in antiretroviral-naïve patients: a prospective study. *AIDS* **15**:647–650.
62. Wang, J., S. J. Smerdon, J. Jager, L. A. Kohlstaedt, P. A. Rice, J. M. Friedman, and T. A. Steitz. 1994. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. *Proc. Natl. Acad. Sci. USA* **91**:7242–7246.
63. Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, and B. H. Hahn. 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**:117–122.
64. Wrobel, J. A., S. F. Chao, M. J. Conrad, J. D. Merker, R. Swanstrom, G. J. Pielak, and C. A. Hutchison III. 1998. A genetic approach for identifying critical residues in the fingers and palm subdomains of HIV-1 reverse transcriptase. *Proc. Natl. Acad. Sci. USA* **95**:638–645.
65. Wu, T. D., C. A. Schiffer, M. J. Gonzales, J. Taylor, R. Kantor, S. Chou, D. Israelski, A. R. Zolopa, W. J. Fessel, and R. W. Shafer. 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. *J. Virol.* **77**:4836–4847.
66. Yang, W., J. P. Bielawski, and Z. Yang. 2003. Widespread adaptive evolution in the human immunodeficiency virus type 1 genome. *J. Mol. Evol.* **57**:212–221.
67. Yerly, S., L. Kaiser, E. Race, J. P. Bru, F. Clavel, and L. Perrin. 1999. Transmission of antiretroviral-drug-resistant HIV-1 variants. *Lancet* **354**:729–733.